

REGRESSION CLUSTERING FOR COMPLEX REAL-WORLD DATA

ABSTRACT OF THE INVENTION

A method and system for determining regression functions from a computer data input using K -Harmonic Means (KHM) regression clustering (RC) and comprising the steps of: (1) selecting K regression functions f_1, \dots, f_K ; (2) associating an i -th data point from the dataset with a k -th regression function using a soft membership function; (3) providing a weighting to each data point using a weighting function to determine the data point's participation in calculating a residue error; (4) calculating the residue error between the weighted i -th data point and its associated regression function; and, (5) iterating to minimize the total residue error. Such can be applied in data mining, economics prediction tools, marketing campaigns, device calibrations, visual image segmentation, and other complex distributions of real-world data.